“On the choice based sample bias in probabilistic bankruptcy prediction”

Kenth Skogsvik (Sweden), Stina Skogsvik (Sweden)

On the choice based sample bias in probabilistic bankruptcy prediction

Abstract

Probabilistic bankruptcy prediction models based on accounting numbers and other financial information are commonly estimated from non-random samples of firms, where the proportion of bankrupt firms is much larger than in most real world situations. This "choice based sample bias" leads to estimated bankruptcy probabilities that are biased. Given that unbiased probabilities are required in risk assessments or discounted cash flow valuation modelling, such probabilities can be severely misleading. The purpose of the paper is to analyze this bias in probabilistic bankruptcy prediction models (typically probit/logit analysis), and to investigate whether it can be mitigated without having to resort to cumbersome model re-estimations. The authors show that there is a clear-cut linkage between sample based probabilities and the corresponding unbiased probabilities. Also, the authors show that sample based probabilities can be calibrated for the choice based sample bias, provided that randomly selected firms from the sub-populations of bankrupt and survival firms are used in the estimation of a prediction model. Non-calibrated bankruptcy probabilities are commonplace in previous empirical research, implying that reported misclassification errors and/or misclassification costs can be more or less misleading. Observed regularities in previous studies are in line with the presented analyses, demonstrating a need for a more insightful treatment of this bias in future research.

Keywords: bankruptcy prediction, business failure, choice based sampling, logit, probabilistic prediction, probit.
JEL Classification: G33, C53.

Introduction

A considerable number of empirical studies on the association between financial statement numbers and firm bankruptcies have been made for industrial, retailing and financial firms over the years[1]. Recently, researchers have taken a particular interest in insurer insolvency prediction[2]. Various statistical techniques have been used to explore this association, ranging from crude applications of regression analysis to more sophisticated variants of probit/logit analysis[3]. It has been observed that the statistical assumptions of regression analysis and discriminant analysis typically are not well fulfilled in the context of bankruptcy prediction[4]. The methods are also somewhat awkward since they do not directly provide estimates of bankruptcy probabilities. In this regard probit/logit analysis is better, as this method implies a probabilistic association between the independent variables (e.g. accounting numbers) and the outcome variable (e.g. "bankruptcy" versus "non-bankruptcy")[5].

Bankruptcy probabilities constitute important parameters in many decision contexts. The relevance of such probabilities in discounted cash flow bond and equity valuation is illustrated in Shaffer (2004) and Skogsvik (2006). Valuation models involving "expected values" presume that it is possible to assess *unbiased* probabilities in the sense that the probabilities are representative for the population of firms. Unbiased bankruptcy probabilities are rarely directly observable, but have to be estimated. In the context of bankruptcy prediction, probit/logit analysis then has appeared to be particularly useful (cf. Ohlson, 1980). However, the issue requires careful attention to the distorting impact of non-random sampling in the estimation of such models.

Since bankruptcies tend to occur rather infrequently, prediction models have in general been estimated from non-random samples of bankrupt and survival firms. The proportion of bankrupt firms in the sample has then typically been much larger than the fraction of such firms in the population. Often a "matched-pairs" design has been used, implying a

The comments of Peter Jennergren, Beate Juettner-Nauroth, Erik Ruist and seminar participants at the Center for Financial Analysis and Managerial Economics in Accounting at Stockholm School of Economics, are gratefully appreciated. The authors are also grateful for financial support from the Torsten and Ragnar Söderberg Foundation and FPG Pensionsgaranti AB.

[1] Early attempts go back to Smith & Winakor (1935) and Merwin (1942), but research of this kind has flourished from the mid 1960s and onwards, cf. Beaver (1966), Altman (1968), Edmister (1972), Altman (1973), Blum (1974), Altman et al. (1977), Ketz (1978), Ohlson (1980), Mensah (1983), Zavgren (1985), Skogsvik (1990), and more lately, Dewaelheyns and Van Hulle (2006), and Pereira Leal and Machado-Santos (2007).

[2] Cf. for example, Cummins et al. (1997), Carson and Hoyt (2000), and Brockett et al. (2006).

[3] Cf. Jones (1987), Jones and Henscher (2004), and Bellovary et al. (2007).

[4] Inconsistent with the assumptions of regression analysis, the dependent variable is dichotomous in a prediction of firm bankruptcy. Multivariate discriminant analysis presumes that the independent variables are multivariate normally distributed in the sub-populations of bankrupt and survival firms, and that the variance/covariance matrices of the independent variables are the same for both sub-populations. As observed in Foster (1986, pp. 107-111) and Skogsvik (1987, pp. 210-214), financial ratios typically do not fulfil these assumptions.

[5] In recent years, other statistical approaches have been suggested in the context of bankruptcy prediction (cf. Shumway, 2001; and Jones & Henscher, 2004). However, standard logit/probit is still the predominant method (Jones & Henscher, 2004).

sample proportion of bankrupt firms of 0.50[1]. The sample proportion of bankrupt firms being exaggerated as compared to the population of firms is the root of the "choice based sample bias", leading to more or less biased probabilities in standard probit/logit models[2].

There are statistical techniques that generate unbiased parameter estimates in probabilistic models, even in the presence of "choice based" sample proportions. These techniques require that the estimation procedure is calibrated for some *a priori* probability of bankruptcy. As long as an estimated prediction model is used in contexts where the *a priori* bankruptcy probability is the same as the proportion of bankrupt firms in the sample, the estimated probabilities will be appropriate. However, if the *a priori* probability in some context does not correspond to the proportion of bankrupt firms in the sample, the estimated probabilities will be biased. Changes in the *a priori* probability of bankruptcy can be due to, for example, variations in firm characteristics, business cycle effects or regulatory changes. A cumbersome way of dealing with this problem would be to re-estimate the prediction model with a new *a priori* probability. However, for a user of some prediction model this might not be a viable alternative – not having access to the original empirical data would obviously be an effective impediment.

The main purpose of this article is to put forward an adjustment formula that will allow users of probabilistic prediction models to eliminate the impact of the choice based sample bias on bankruptcy probabilities. The paper provides guidelines for the use of probabilistic prediction models in out-of-sample contexts, potentially valuable for both academics and practitioners. Additionally, empirical consequences of not making any adjustments to model-based probabilities are addressed, with references to previous research.

The outline of the article is as follows. The choice based sample bias of bankruptcy probabilities is analyzed in section 1. In section 2, the ranking characteristics of biased probabilities are evaluated, and in section 3 the calibration of biased probabilities is analyzed. Section 4 provides guidelines for choosing probability cut-off values when evaluating the classification accuracy of probabilistic models. Implications for the evaluation of the prediction performance of bankruptcy prediction models is discussed in section 5. The last section concludes the paper.

1. The impact of the choice based sample bias on assessed probabilities

Probabilistic bankruptcy prediction models have commonly been estimated from non-random samples in previous research. Given that standard (unweighted) statistical techniques have been used, estimated coefficients have then been affected by the chosen sample proportions. In order to analyze this effect, we use the following notation: $\pi(t)$ is the proportion of bankrupt firms year t in the population of firms, i.e. the *a priori* probability of bankruptcy year t; $p^{(\pi)}_{j,fail(t)}$ is the *unbiased probability* of bankruptcy (consistent with the *a priori* probability of bankruptcy in the population) year t for firm j, conditioned on firm survival at the end of year t-1; *prop* is the proportion of bankrupt firms in the estimation sample; $p^{(prop)}_{j,fail(t)}$ is the *sample based probability* of bankruptcy (consistent with the proportion of bankrupt firms in the estimation sample) year t for firm j, conditioned on firm survival at the end of year t-1; and $\{X_{j,t\text{-}1}\}$ is the set of financial descriptors for firm j, observable at time t-1.

In a decision context, the idea is that a decision maker is armed with some (previously estimated) bankruptcy prediction model, and that:

- Values of the descriptor variables for some firm $\{X_{j,t\text{-}1}\}$ (including accounting numbers as indicators of profitability, interest cost, financial leverage, etc.)[3] are measured.
- Based on the probabilistic prediction model, a probability of firm bankruptcy $\left(p^{(prop)}_{j,fail(t)}\right)$ is calculated over some forecast horizon.

In order to simplify the notation, let henceforth the indices j and t be suppressed. Recognizing that both sample based and unbiased probabilities are conditioned on the set of descriptor variables, we can then write:

$$p^{(prop)}_{fail} = p\big(fail \,|\{X\}\big)^{(prop)},$$

$$p^{(\pi)}_{fail} = p\big(fail \,|\{X\}\big)^{(\pi)}.$$

Assuming that $0 < \pi < 1.00$, the unbiased bankruptcy probability $p^{(\pi)}_{fail}$ can be analyzed in accordance with Bayes theorem as follows[4]:

[1] In a survey of failure prediction studies, Zmijewski (1984) observed that a "matched-pairs" design had been used in about 70% of previous studies. Similarly, "matched-pairs" sampling has been predominant in insurer bankruptcy studies (Carson and Hoyt, 2000).

[2] Ohlson (1980) constitutes an exception from the "matched-pairs" sampling procedure, as Ohlson's sample included 105 bankrupt and 2058 survival US industrial firms (i.e. the proportion of failure firms was less than 5%.) One might hence expect that Ohlson´s prediction model became comparatively representative for the population of US industrial firms, which might have contributed to the fairly robust performance of this model over time (cf. Begley et al., 1997; and Boritz et al., 2007).

[3] Cf. for example, Skogsvik (1990, p. 145 and pp. 155-157).

[4] Cf. for example, Chou (1984, p. 411).

$$p_{fail}^{(\pi)} = p\left(fail\,|\{X\}\right)^{(\pi)} = \frac{\pi \cdot p\left(\{X\}|\,fail\right)^{(\pi)}}{\pi \cdot p\left(\{X\}|\,fail\right)^{(\pi)} + (1-\pi)\cdot p\left(\{X\}|\,surv\right)^{(\pi)}} =$$

$$= \left[1 + \frac{(1-\pi)}{\pi}\cdot\frac{p\left(\{X\}|\,surv\right)^{(\pi)}}{p\left(\{X\}|\,fail\right)^{(\pi)}}\right]^{-1}, \quad (1)$$

where $p\left(\{X\}|\,fail\right)^{(\pi)}$ is the unbiased probability of observing $\{X\}$ (at time t) conditioned on firm bankruptcy in year (t+1); and $p\left(\{X\}|\,surv\right)^{(\pi)}$ is the unbiased probability of observing $\{X\}$ (at time t) conditioned on firm survival in year (t+1).

The sample based probability $p_{fail}^{(prop)}$ is affected by the proportion of bankrupt firms in the estimation sample, where *prop* can be viewed as the "*a priori*" fraction of bankrupt firms in the sample. Assuming that 0 < *prop* < 1.00, this probability can be analyzed in the same manner as the unbiased probability $p_{fail}^{(\pi)}$:

$$p_{fail}^{(prop)} = p\left(fail\,|\{X\}\right)^{(prop)} =$$

$$= \left[1 + \frac{(1-prop)}{prop}\cdot\frac{p\left(\{X\}|surv\right)^{(prop)}}{p\left(\{X\}|fail\right)^{(prop)}}\right]^{-1}, \quad (2)$$

where $p\left(\{X\}|\,fail\right)^{(prop)}$ is the sample probability of observing $\{X\}$ (at time t) conditioned on firm bankruptcy the following year (t+1); and $p\left(\{X\}|\,surv\right)^{(prop)}$ is the sample probability of observing $\{X\}$ (at time t) conditioned on firm survival the following year (t + 1).

We now presume that the sample of bankrupt firms constitutes a random drawing from the sub-population of bankrupt firms and the sample of survival firms constitutes a random drawing from the sub-population of survival firms, in the sense that $p\left(\{X\}|\,fail\right)^{(prop)} = p\left(\{X\}|\,fail\right)^{(\pi)}$ and $p\left(\{X\}|\,surv\right)^{(prop)} = p\left(\{X\}|\,surv\right)^{(\pi)}$. Given that both $p\left(\{X\}|\,fail\right)^{(\pi)}$ and $p\left(\{X\}|\,fail\right)^{(prop)}$ are positive, this means $p\left(\{X\}|\,surv\right)^{(\pi)}/p\left(\{X\}|\,fail\right)^{(\pi)} = p\left(\{X\}|\,surv\right)^{(prop)}/p\left(\{X\}|\,fail\right)^{(prop)}$. Let this *odds ratio* be denoted $Or(\{X\})$.

The odds ratio for the population of firms can be solved through a rewriting of equation (1), i.e.:

$$Or\left(\{X\}\right) = \left[\frac{1}{p\left(fail\,|\,\{X\}\right)^{(\pi)}} - 1\right]\cdot\frac{\pi}{(1-\pi)}. \quad (3)$$

Inserting the above solution for $Or(\{X\})$ in equation (2), and recognizing the equalities

$p\left(fail\,|\{X\}\right)^{(\pi)} = p_{fail}^{(\pi)}$ and $p\left(fail|\{X\}\right)^{(prop)} = p_{fail}^{(prop)}$,

we get:

$$p_{fail}^{(prop)} = \left[1 + \left(\frac{\pi}{1-\pi}\right)\cdot\left(\frac{1-prop}{prop}\right)\cdot\left(\frac{1-p_{fail}^{(\pi)}}{p_{fail}^{(\pi)}}\right)\right]^{-1}. \quad (4)$$

The sample based bankruptcy probability is hence a function of the unbiased probability $p_{fail}^{(\pi)}$, the fraction of bankrupt firms in the population (π), and the proportion of bankrupt firms in the estimation sample (*prop*). As expected, equation (4) shows that $p_{fail}^{(prop)} = p_{fail}^{(\pi)}$ if the proportion of bankrupt firms in the sample is equal to the *a priori* bankruptcy probability. However, if *prop* = 0.5 (as in matched pairs sampling) and π = 0.02, the value of $p_{fail}^{(prop)}$ would be $\left[\frac{p_{fail}^{(\pi)}}{0.0204 + 0.9796\cdot p_{fail}^{(\pi)}}\right]$, meaning that the sample based probability would be exaggerated as long as $0 < p_{fail}^{(\pi)} < 1.00$. For example, setting $p_{fail}^{(\pi)}$ alternatively to 0.01, 0.02 or 0.10, the sample based probability $p_{fail}^{(prop)}$ in a matched-pairs sampling design would be equal to 0.33, 0.50 and 0.84, respectively.

In order to better understand the linkage between sample based probabilities and the proportion of bankrupt firms in the estimation sample, we can calculate the derivative of equation (4) with respect to *prop*:

$$\frac{\partial\left(p_{fail}^{(prop)}\right)}{\partial(prop)} = \frac{\pi\cdot(1-\pi)\cdot p_{fail}^{(\pi)}\cdot\left(1-p_{fail}^{(\pi)}\right)}{\left[\pi\cdot\left(p_{fail}^{(\pi)} + prop - 1\right) - prop\cdot p_{fail}^{(\pi)}\right]^2}. \quad (5)$$

Limiting the analysis to settings where $0 < \pi < 1.00$, $0 < prop < 1.00$ and $0 < p_{fail}^{(\pi)} < 1.00$, the RHS of (5) is positive, in turn meaning that $p_{fail}^{(prop)}$ is positively affected by the sample proportion of bankrupt firms. Alternatively, given that $prop > \pi$, equation (4) and (5) imply that the bias of $p_{fail}^{(prop)}$ is positive and increasing in the proportion of bankrupt firms in the estimation sample.

2. Ranking characteristics of sample based probabilities

In accordance with equation (4) above, a sample based bankruptcy probability ($p_{fail}^{(prop)}$) constitutes a biased assessment of the probability $p_{fail}^{(\pi)}$ when $prop \neq \pi$. However, in some decision contexts, estimated probabilities are only used for classifying firms as bankrupt or non-bankrupt entities. The focus will then be on the ranking of firms based on the sample based probabilities, in combination with a chosen probability cut-off value. It is thus interesting to know whether a sample based probability $p_{fail}^{(prop)}$, being more or less biased in relation to $p_{fail}^{(\pi)}$, nevertheless might provide a correct ranking of firms. That is, reintroducing the company index and letting *j* and *j'* denote two different companies − if $p_{fail,j}^{(\pi)} > p_{fail,j'}^{(\pi)}$ will then $p_{fail,j}^{(prop)} > p_{fail,j'}^{(prop)}$?

In order to answer this question, we calculate the derivative of equation (4) with respect to the unbiased probability $p_{fail}^{(\pi)}$:

$$\frac{\partial\left(p_{fail}^{(prop)}\right)}{\partial\left(p_{fail}^{(\pi)}\right)} = \frac{\pi\cdot(1-\pi)\cdot prop\cdot(1-prop)}{\left[\pi\cdot\left(p_{fail}^{(\pi)}+prop-1\right)-prop\cdot p_{fail}^{(\pi)}\right]^2} > 0. \quad (6)$$

Equation (6) shows that there is a positive relationship between $p_{fail}^{(prop)}$ and $p_{fail}^{(\pi)}$, implying that a sample based probability is a *consistent* indicator of bankruptcy risk in the following sense:

Proposition: Let $p_{fail,j}^{(\pi)}$ denote the unbiased bankruptcy probability and $p_{fail,j}^{(prop)}$ the sample based bankruptcy probability of firm *j*. Then it holds:

- if $p_{fail,j}^{(\pi)} > p_{fail,j'}^{(\pi)}$, then $p_{fail,j}^{(prop)} > p_{fail,j'}^{(prop)}$;
- if $p_{fail,j}^{(\pi)} = p_{fail,j'}^{(\pi)}$, then $p_{fail,j}^{(prop)} = p_{fail,j'}^{(prop)}$;
- if $p_{fail,j}^{(\pi)} < p_{fail,j'}^{(\pi)}$, then $p_{fail,j}^{(prop)} < p_{fail,j'}^{(prop)}$.

Given that a probabilistic prediction model has been estimated with a sample proportion of bankrupt firms $prop \neq \pi$, the probability $p_{fail}^{(prop)}$ is biased but according to the Proposition the ranking of firms with the biased and the unbiased probabilities will nevertheless be the same. Hence the choice based sample bias in probabilistic bankruptcy prediction does not affect the reliability of $p_{fail}^{(prop)}$, only its ability to correctly depict the unbiased probability $p_{fail}^{(\pi)}$.

3. Estimating unbiased bankruptcy probabilities

In the previous section it was shown that even if the probability $p_{fail}^{(prop)}$ is biased, it still provides a correct ranking of firms. In a typical bond or equity valuation problem however, a decision maker needs to transform sample based bankruptcy probabilities into their unbiased counterparts. This section shows how to transform such biased probabilities into unbiased probabilities.

According to equation (4), a sample based probability can be written as a function of the unbiased probability, the fraction of bankrupt firms in the population, and the proportion of bankrupt firms in the estimation sample. Our focus is now on estimating the probability $p_{fail}^{(\pi)}$. Rewriting equation (4) gives the following adjustment formula for how to calibrate the sample based probability:

$$p_{fail}^{(adj)} = \left[1+\left(\frac{1-\pi}{\pi}\right)\cdot\left(\frac{prop}{1-prop}\right)\cdot\left(\frac{1-p_{fail}^{(prop)}}{p_{fail}^{(prop)}}\right)\right]^{-1}, \quad (7)$$

where $p_{fail}^{(adj)}$ is the sample based probability of bankruptcy year *t* (for firm *j*) conditioned on firm survival at the end of year *t*-1, calibrated for the fraction of failure companies in the population.

Equation (7) shows how an unbiased probability $p_{fail}^{(adj)}$ can be calculated as a function of the fraction of bankrupt firms in the population (π), the proportion of bankrupt firms in the estimation sample (*prop*) and the biased bankruptcy probability ($p_{fail}^{(prop)}$). For example, if the bankruptcy frequency in the population is 0.02, the sample proportion of bankrupt firms is 0.50, and an estimated prediction model generates $p_{fail}^{(prop)}$ = 0.60, the calibrated bankruptcy risk would be

$$p_{fail}^{(adj)} = \left[1+\frac{(1-0.02)}{0.02}\cdot\frac{0.50}{(1-0.50)}\cdot\frac{(1-0.60)}{0.60}\right]^{-1} = 0.03.$$[1]

Note that $p_{fail}^{(adj)}$ in equation (7) constitutes an *estimate* of the unbiased probability $p_{fail}^{(\pi)}$. As standard probit/logit techniques do not provide any sampling errors associated with $p_{fail}^{(prop)}$, it is hard to make a precise statement on the sampling characteristics of $p_{fail}^{(adj)}$. However (as stated in section 1), a necessary condition for $p_{fail}^{(adj)}$ to be an unbiased estimator of $p_{fail}^{(\pi)}$ is that the samples of bankrupt and survival firms constitute random drawings from the sub-populations of bankrupt and survival firms, respectively.

4. Implications for the use of probabilistic prediction models

A couple of methodological consequences of the choice based sample bias will be addressed in this section. The first issue is concerned with the

[1] Evidently, if there is no choice based sample bias, then $prop = \pi$ and $p_{fail}^{(adj)} = p_{fail}^{(prop)}$ in equation (7).

magnitude of estimated coefficients in standard probit/logit models and the second issue deals with the classification accuracy of estimated prediction models.

Regarding the magnitude of the estimated coefficients, equation (4) shows that $p_{fail}^{(prop)}$ will be positively biased when *prop* > π. When the proportion of bankrupt firms in the estimation sample is larger than the fraction of such firms in the population (as typically has been the case in previous empirical research), the estimated coefficients can hence be expected to be "exaggerated". However, as more carefully discussed in Manski & Lerman (1977), it is difficult to more precisely specify the bias of the coefficients of the independent variables. Empirical tests in Zmijewski (1984) and Bergström et al. (1999) show that significance tests of the coefficients in the main appear to be unaffected by variations in *prop*, at least as long as there are 40 or more bankrupt firms in the estimation sample.

Concerning the classification accuracy of bankruptcy prediction models, tests of this kind involves the choice of a cut-off value $\overline{p}_{fail}$ such that firms with $p_{fail}^{(prop)} > (<) \overline{p}_{fail}$ are classified as bankrupt (survival) firms. Defining "error type I" as an erroneous classification of a bankrupt firm and "error type II" as an erroneous classification of a survival firm, it is easily recognized that the choice of $\overline{p}_{fail}$ involves a trade-off between the size of type I and type II errors[1].

Regarding the choice of $\overline{p}_{fail}$, basically two approaches have been used in previous research − an "empirical" and an "analytical" approach. According to the former approach, a cut-off value, here denoted $\overline{p}_{fail}^{(emp)}$, is determined empirically as the cut-off probability associated with the lowest "average error rate", or "average error cost", for the estimation sample. As regards the measurement of the average error rate, ($rate(\overline{e})$), the following definitions have been used in previous research:

$$rate(\overline{e}) = [\, rate(e_{\mathrm{I}}) + rate(e_{\mathrm{II}})]/2, \tag{8a}$$

$$rate(\overline{e})' = prop \cdot rate(e_{\mathrm{I}}) + (1 - prop) \cdot rate(e_{\mathrm{II}}), \tag{8b}$$

$$rate(\overline{e})'' = \pi \cdot rate(e_{\mathrm{I}}) + (1 - \pi) \cdot rate(e_{\mathrm{II}}), \tag{8c}$$

where $rate(e_{\mathrm{I}})$ is error rate type I, i.e. the number of errors type I in relation to the number of bankrupt firms in the sample; and $rate(e_{\mathrm{II}})$ is error rate type II, i.e. the number of errors type II in relation to the number of survival firms in the sample.

The error rate in equation (8a) is simply the arithmetic average of error rates type I and type II, while the error rates in (8b) and (8c) are functions of the relative frequency of bankrupt firms in the estimation sample and the population, respectively. Note that, if 0.5 = *prop* = π the average error rates in equations (8a) to (8c) coincide, in turn meaning that the corresponding cut-off values $\overline{p}_{fail}^{(emp)}$ will be the same.

Alternatively, an empirical cut-off value $\overline{p}_{fail}^{(emp)}$ can be determined as the cut-off probability that minimizes the average error cost, determined as:

$$\overline{cost} = w_1 \cdot cost_{\mathrm{I}} + w_2 \cdot cost_{\mathrm{II}}, \tag{9}$$

where $cost_{\mathrm{I}}$ is the cost associated with a classification error type I; $cost_{\mathrm{II}}$ is the cost associated with a classification error type II; w_1 is the weight of error cost type I; and w_2 is the weight of error cost type II.

In previous research, the weights w_1 and w_2 have commonly been specified as a function of the fraction of bankrupt firms in the population and empirically estimated values of $rate(e_{\mathrm{I}})$ and $rate(e_{\mathrm{II}})$, giving an average error cost equal to:

$$\overline{cost}^{(\pi)} = [\pi \cdot rate(e_{\mathrm{I}})] \cdot cost_{\mathrm{I}} + \\ + [(1-\pi) \cdot rate(e_{II})] \cdot cost_{\mathrm{II}}. \tag{10}$$

Probability cut-off values based on equations (8a), (8b), (8c) and (9) are affected by the choice based sample bias in the same way as $p_{fail}^{(prop)}$, i.e. $\overline{p}_{fail}^{(emp)}$ will be positively (negatively) biased if *prop* > π (*prop* < π). In principle, this bias is harmless as long as the cut-off values are used to evaluate correspondingly biased values of $p_{fail}^{(prop)}$. Since rankings based on $p_{fail}^{(prop)}$ and $p_{fail}^{(\pi)}$ coincide, there will always exist a biased cut-off value $\overline{p}_{fail}^{(emp)}$ that generates the same average error rate, or error cost, as an unbiased probability cut-off value together with unbiased probabilities $p_{fail}^{(adj)}$[2].

The analytical approach for determining a probability cut-off value is derived from some decision context where unbiased probabilities are presumed. In previous research this choice has often been guided by a simple trade-off between expected error costs, calculated as:

- Expected error cost of "survival" classification: $p_{fail}^{(\pi)} \cdot cost_{\mathrm{I}}$.
- Expected error cost of "bankruptcy" classify-cation: $\left(1 - p_{fail}^{(\pi)}\right) \cdot cost_{\mathrm{II}}$.

[1] As one extreme, if $\overline{p}_{fail}$ = 0 there will be no errors type I but all survival firms will be classified as "bankrupt firms" (errors type II). On the other hand, if $\overline{p}_{fail}$ = 1 there will be no errors type II, but all bankrupt firms will be classified as "survival firms" (errors type I).

[2] Cf. for example, Skogsvik (1990, pp. 149-150), and in the context of probabilistic predictions of firm profitability, Skogsvik (2008, pp. 803-804).

Decision rule. A "bankruptcy" classification is made when $p_{fail}^{(\pi)} \cdot cost_{\mathrm{I}} > \left(1 - p_{fail}^{(\pi)}\right) \cdot cost_{\mathrm{II}}$; otherwise a "survival" classification is made.

The probability cut-off value, $\bar{p}^{*}_{fail}$, implied by this decision rule can easily be solved[1]:

$$\bar{p}^{*}_{fail} = \frac{1}{1 + (cost_{\mathrm{I}} / cost_{II})}. \tag{11}$$

In contrast to empirically determined probability cut-off values, analytically derived cut-off values only make sense when unbiased bankruptcy probabilities are available. As noted previously, $p_{fail}^{(prop)}$ is biased when $prop \neq \pi$ and such probabilities should not be used with analytically derived cut-off values. One way to deal with this issue is to first transform the sample based probabilities $p_{fail}^{(prop)}$ into unbiased probabilities $p_{fail}^{(adj)}$ in accordance with equation (7), and then use an analytically derived probability cut-off value. Alternatively, the problem could be handled by the adjustment equation (4), which transforms unbiased probabilities into biased probabilities, i.e.:

$$\bar{p}^{*(prop)}_{fail} = \left[1 + \left(\frac{\pi}{1-\pi}\right) \cdot \left(\frac{1-prop}{prop}\right) \cdot \left(\frac{1-\bar{p}^{*}_{fail}}{\bar{p}^{*}_{fail}}\right)\right]^{-1}. \tag{12}$$

$\bar{p}^{*(prop)}_{fail}$ here constitutes the adjusted analytical cut-off probability $\bar{p}^{*}_{fail}$, to be used with *unadjusted* bankruptcy probabilities $p_{fail}^{(prop)}$.

5. Implications for reported classification/ prediction results in previous research

A vast number of bankruptcy prediction models have been estimated and tested over the years, typically without any consideration of the importance of the choice based sample bias. An interesting issue is thus to what extent this negligence has misguided or distorted the evaluation of these models.

Addressing various methodological issues related to standard probabilistic bankruptcy prediction modeling, the classification accuracy − measured as $rate(\bar{e})'$ in equation (8b) above − for different proportions of bankrupt firms in the estimation sample was calculated in Zmijewski (1984). In all classification tests a cut-off probability of 0.50 was used, presumably viewed as an analytical cut-off value based on a symmetric loss function (i.e. $cost_{\mathrm{I}}/cost_{\mathrm{II}}$ = 1.00)[2]. With regard to the observed results, Zmijewski stated[3]:

"The results… generally indicate the existence of a bias and the overclassification of bankrupt firms when using unweighted probit" (Zmijewski, 1984, p. 72).

Zmijewski's observation is not surprising. As the bias of $p_{fail}^{(prop)}$ is positively related to the sample proportion of bankrupt firms, there will be more firms with $p_{fail}^{(prop)}$ being larger than $\bar{p}^{*}_{fail}$ = 0.50 when this proportion is high. Hence, a lower fraction of misclassified bankrupt firms − and a higher fraction of misclassified survival firms − trivially follows.

The prediction accuracy was evaluated by Zmijewski with a holdout sample including 41 bankrupt and 800 survival firms, implying a fraction of bankrupt firms in the holdout sample of 0.049. The impact of the choice based sample bias on $p_{fail}^{(prop)}$ was clearly observed for this sample. When the proportion of bankrupt firms in the estimation sample was 0.50, the average value of $p_{fail}^{(prop)}$ was 0.19 in the holdout sample. When the proportion of bankrupt firms in the estimation sample was reduced − to 0.286, 0.167, 0.091 and 0.048 − the average value of $p_{fail}^{(prop)}$ decreased − to 0.11, 0.09, 0.07, 0.06 and 0.05, respectively[4]. Consistent with our analysis in section 1 above, average values of $p_{fail}^{(prop)}$ were larger than π when $prop > 0.048$ and the average values of the sample based probabilities decreased as *prop* decreased. It is particularly worth noting that, when the proportion of bankrupt firms in the estimation sample was about the same as the fraction of such firms in the holdout sample (0.048 ≈ 0.049), the average value of $p_{fail}^{(prop)}$ (= 0.05) was very close to the "*a priori* probability" of bankruptcy in the holdout sample (0.049). This illustrates that $p_{fail}^{(prop)}$ is an unbiased estimate of $p_{fail}^{(\pi)}$ when $prop = \pi$ (in line with equation (4) above). Hence, our analysis helps to clarify the empirical observations in Zmijewski (1984).

Zmijewski also calculated weighted average error rates in accordance with equation (8c), setting the probability cut-off value to 0.50 for the estimated probit models. The reported results were as follows:

"… the bankrupt firm correlation is positive … indicating an overclassification bias; the nonbankrupt

[1] The probability cut-off value is implied in equation (11) in the sense that it is rational for a risk-neutral decision maker to classify a firm as "bankrupt" if $p_{fail}^{(\pi)} > \bar{p}^{*}_{fail}$ and "non-bankrupt" if $p_{fail}^{(\pi)} < \bar{p}^{*}_{fail}$.

[2] Cf. note 16, p. 72, in Zmijewski (1984).

[3] "Unweighted probit" in the quotation refers to an application of probit analysis where no adjustments are made to handle the choice based sample bias.

[4] From Table 5 in Zmijewski (1984, p. 71).

firm correlation is negative … indicating an under-classification bias; and the overall correlation is negative, indicating that correct prediction rates increase when parameters which are less biased are used" (Zmijewski, 1984, p. 73).

The overclassification bias for the bankrupt firms and the underclassification bias for the non-bankrupt firms are both consistent with the choice based sample bias effect on $p_{fail}^{(prop)}$, as noted above. However, claiming that the average error rate decreases as the model parameters are more unbiased, is more doubtful. As the chosen cut-off probability (0.50) was distinctively different from the proportion of bankrupt firms in the holdout sample (0.049), the cut-off value presumably was expected to be optimal in a decision context where $cost_I/cost_{II}$ = 1.00. Minimizing the average error cost in equation (10) when $cost_I/cost_{II}$ = 1.00 is equivalent to minimizing the average error cost calculated as

$$\overline{cost}^{(\pi)} = \pi \cdot rate(e_I) \cdot 1.0 + (1-\pi) \cdot rate(e_{II}) \cdot 1.0,$$

i.e. in *this particular instance* the same as minimizing the weighted average error rate in equation (8c). As argued previously, analytically derived cut-off probabilities are only consistent with unbiased probabilities, i.e. $p_{fail}^{(prop)}$ based on estimation samples for which $prop = \pi$ or estimates of unbiased probabilities ($p_{fail}^{(adj)}$). Since the fraction of bankrupt firms in the holdout sample was 0.049 in Zmijewski's tests of prediction performance, the estimation sample with the lowest proportion of failure companies (*prop* = 0,048) should have generated the most unbiased values of $p_{fail}^{(prop)}$. It is then only to be expected that the lowest average error rate in equation (8c) should be observed for this estimation sample. However, about the same average error rate should also have been possible to observe if the biased probabilities $p_{fail}^{(prop)}$ from the other estimation samples had been calibrated in accordance with equation (7) above.

Also, the analyses in Zmijewski (1984) fail to recognize that an analytical probability cut-off value $\overline{p}_{fail}^{*}$ is linked to a specific decision context and that the evaluation of prediction models should be based on the goal function of that particular context. Setting $\overline{p}_{fail}^{*}$ = 0.50 in Zmijewski (1984), the average error cost according to equation (10) *happens* to coincide with the average error rate in equation (8c). However, if − as one typically would assume − $cost_I/cost_{II}$ > 1,00, the cut-off probability should be less than 0.50 and the average error rate in equation (8c) would *not* have constituted a valid indicator of prediction performance. In such cases, estimated average *error costs* in accordance with equation (10) should always be used to assess the prediction performance of probabilistic models[1].

Conclusions

The purpose of the article has been to enhance the usefulness of (unweighted) probabilistic bankruptcy prediction models. Specifically, problems associated with the choice based sample bias have been addressed, as previous research is vague and sometimes even misleading on this issue. Future empirical research can benefit from the provided guidelines of how to handle the choice based sample bias in this type of prediction modeling.

Bankruptcy probabilities have commonly been estimated in probabilistic statistical models, as, for example, in Ohlson (1980), Zavgren (1985), Skogsvik (1990) and more lately, Dewaelheyns & Van Hulle (2006) and Pereira Leal & Machado-Santos (2007). A consequence of the choice based sample bias is that estimated probabilities in standard probit/logit models are not representative for the population, if the proportion of bankrupt firms in the estimation sample differs from the corresponding fraction in the population. In previous empirical research, the proportion of bankrupt firms in estimation samples has commonly been distinctively larger than the corresponding population fraction.

We have shown that there is a specific linkage between sample based and unbiased (population based) probabilities. A sample based probability $p_{fail}^{(prop)}$ is a function of the unbiased probability $p_{fail}^{(\pi)}$, the proportion of bankrupt firms in the estimation sample, and the fraction of such firms in the population. Characterizing this linkage it was found that $p_{fail}^{(prop)} > p_{fail}^{(\pi)}$ if $prop > \pi$ and vice versa, but that the ranking of firms based on $p_{fail}^{(prop)}$ or $p_{fail}^{(\pi)}$ is the same. Having specified the linkage between $p_{fail}^{(prop)}$ and $p_{fail}^{(\pi)}$, the choice based sample bias of $p_{fail}^{(prop)}$ can be disentangled to get a calibrated probability. There is hence no need to re-estimate probit/logit models only because the sample proportion of bankrupt firms is non-representative. Rather, it is important to collect sufficiently large samples of randomly selected bankrupt and survival firms.

In sum, there are alternative approaches for handling the choice-based sample bias depending on the decision context. In situations where bankruptcy

[1] In line with, for example, the methodology outlined in Skogsvik (1990, p. 150, in particular footnote 22).

probabilities are only used for classifying companies as bankrupt or survival firms, the choice-based sample bias is (in expectation) unproblematic as long as the probability cut-off value is empirically assessed. This approach has for example been used in Ohlson (1980), Zavgren (1985), and Skogsvik (1990). As a contrast, if the probability cut-off value is analytically derived, either the cut-off value itself has to be calibrated in order to be applicable to biased firm-specific probabilities, or the firm-specific bankruptcy probabilities have to be transformed into unbiased probabilities. However, in many situations unbiased probabilities are required, e.g. in financial risk management or in equity or bond valuation problems. Biased probabilities should then be transformed into unbiased probabilities by applying the calibration equation (7).

An alternative approach to handle the choice based sample bias would be to estimate bankruptcy prediction models using "weighted" probabilistic statistical techniques, as suggested in Zmijewski (1984, p. 74). Since an *a priori* probability of bankruptcy has to be specified when estimating a model according to these techniques, re-estimations of the prediction model would then be necessary as soon as this *a priori* probability changes (due to e.g. shifting business conditions). Obviously, the approach proposed in the paper is more flexible and less costly, since no re-estimation of the prediction model is needed.

Empirical observations in previous research can be explained by our results. In line with our inferences, average values of $p_{fail}^{(prop)}$ have been found to be larger than $p_{fail}^{(\pi)}$ when $prop > \pi$, and to increase as the proportion of bankrupt firms in the estimation sample goes up. Furthermore, with a probability cut-off value equal to 0.50, an "overclassification" bias of bankrupt firms and an "underclassification" bias of survival firms have been observed for samples where $prop > \pi$. This follows from the bias of $p_{fail}^{(prop)}$ being positive in situations when $prop > \pi$. However, suggestions in previous research that prediction results improve if the choice based sample bias is reduced, are misleading.

References

1. Altman, E.I. (1968). "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", *The Journal of Finance*, No. 4.
2. Altman, E.I. (1973). "Predicting Railroad Bankruptcies in America", *Bell Journal of Economics and Management Science*, Spring.
3. Altman, E.I., Haldeman, R.G. & Narayanan, P. (1977). "Zeta^{TM} Analysis – A New Model to Identify Bankruptcy Risk of Corporations", *Journal of Banking and Finance*, No. 1, June.
4. Beaver, W.H. (1966). "Financial Ratios as Predictors of Failure", *Journal of Accounting Research*, Supplement to Vol. 4.
5. Begley, J., Ming, J. & Watts, S. (1997). "Bankruptcy Classification Errors in the 1980s: An Empirical Analysis of Altman's and Ohlson's Model", *Review of Accounting Studies*, Vol. 1, No. 4.
6. Bellovary, J.L.G., Giacomoni, D.E. & Akers, M.D. (2007). "A Review of Bankruptcy Prediction Studies: 1930 to present", SSRN working papers.
7. Bergström, C., Eisenberg, T., Sundgren, S. & Wells, M.T. (1999). "Methodological issues in bankruptcy prediction", in Gren, B. (editor), *Risk Behaviour and Risk Management in Business Life*, Kluwer Academic Publishers.
8. Blum, M. (1974). "Failing Company Discriminant Analysis", *Journal of Accounting Research*, No. 1.
9. Boritz, J.E., Kennedy, D.B. & Sun, J.Y. (2007). "Predicting Business Failures in Canada", SSRN working paper.
10. Brockett, P.L., Golden, L.L., Jang, J. & Yang, C. (2006). "A Comparison of Neural Network, Statistical Methods, and Variable Choice for Life Insurers' Financial Distress Prediction", *The Journal of Risk and Insurance*, Vol. 73.
11. Carson, J. & Hoyt, R.E. (1995). "Life Insurer Financial Distress: Classification Models and Empirical Evidence", *Journal of Risk and Insurance*, 62.
12. Carson, J. & Hoyt, R. (2000). "Evaluating the Risk of Life Insurer Insolvency: Implications from the US for the European Union", *Journal of Multinational Financial Management*, 10.
13. Chou, Y. (1984). *Statistical Analysis*s (second edition), Holt, Rinehart and Winston.
14. Cummins, J.D., Grace, M.F. & Phillips, R.D. (1999). "Regulatory Solvency Prediction in Property-Liability Insurance: Risk-Based Capital, Audit Ratios, and Cash Flow Simulation", *Journal of Risk and Insurance*, 66 (3).
15. Dewaelheyns, N. & Van Hulle, C. (2006). "Corporate Failure Prediction Modeling: Distorted by Business Groups' Internal Capital Markets?", *Journal of Business Finance & Accounting*, Vol. 33, No. 5-6.
16. Edmister, R.O. (1972). "Test of Financial Ratio Analysis for Small Business Failure Prediction", *Journal of Financial and Quantitative Analysis*, March.
17. Foster, G. (1986). *Financial Statement Analysis* (second edition), Prentice-Hall International.
18. Jones, F. (1987). "Current Techniques in Bankruptcy Prediction", *Journal of Accounting Literature*, 6.
19. Jones, S. & Henscher, D.A. (2004). "Predicting Firm Financial Distress: A Mixed Logit Model", *The Accounting Review*, Vol. 79, No. 4.

20. Ketz, J.E. (1978). "The Effect of General Price-Level Adjustments on the Predictive Ability of Financial Ratios", *Journal of Accounting Research*, supplement to Volume 16.
21. Manski, C.F. & Lerman, S.R. (1977). "The Estimation of Choice Probabilities from Choice Based Samples", *Econometrica*, 45 (8), November.
22. Mensah, V.M. (1983). "The Differential Bankruptcy Predictive Ability of Specific Price Level Adjustments: Some Empirical Evidence", *The Accounting Review*, April.
23. Merwin, C. (1942). *Financing Small Corporations*, National Bureau of Economic Research, New York.
24. Ohlson, J.A. (1980). "Financial Ratios and the Probabilistic Prediction of Bankruptcy", *Journal of Accounting Research*, 18 (1), Spring.
25. Pereira Leal, C. & Machado-Santos, C. (2007). "Insolvency Prediction in the Portuguese Textile Industry", *European Journal of Finance and Banking Research*, Vol. 1, No. 1.
26. Schaffer, S. (2004). "Corporate Failure and Equity Valuation", Working paper, University of Wyoming.
27. Shumway, T. (2001). "Forecasting Bankruptcy More Accurately: A Simple Hazard Model", *Journal of Business*, Vol. 74, No. 1.
28. Skogsvik, K. (1990). "Current Cost Accounting Ratios as Predictors of Business Failure: The Swedish Case", Journal of Business Finance & Accounting, 17 (1), Spring.
29. Skogsvik, K. (2006). "Probabilistic Business Failure Prediction in Discounted Cash Flow Bond and Equity Valuation", SSE/EFI Working Paper Series in Business Administration (SWOBA), 5.
30. Skogsvik, S. (2008). "Financial Statement Information, the Prediction of Book Return on Owners' Equity and Market Efficiency: The Swedish Case", *Journal of Business Finance & Accounting*, 35 (7) & (8).
31. Smith, R.F. & Winakor, A.H. (1935). *Changes in the Financial Structure of Unsuccessful Corporations*, University of Illinois, Bureau of Business Research.
32. Zavgren, C.V. (1985). "Assessing the Vulnerability to Failure of American Industrial Firms: A Logistic Analysis", *Journal of Business Finance and Accounting*, Spring.
33. Zmijewski, M.E. (1984). "Methodological Issues Related to the Estimation of Financial Distress Prediction Models", *Journal of Accounting Research*, 22, Supplement.